NORTH-HOLLAND

A Determinantal Proof of the Craig-Sakamoto Theorem

Ingram Olkin
Department of Statistics
Stanford University
Stanford, California 94305-4065

Submitted by George P. H. Styan

ABSTRACT

The Craig-Sakamoto theorem states that if A and B are symmetric matrices, then (a) $|I - \alpha A - \beta B| = |I - \alpha A||I - \beta B|$ for all α, β if and only if (b) $AB = 0$. There are a number of proofs of this result, the most common based on expansions of the logarithm of (a). The present proof is elementary in that it depends only on determinantal conditions. © 1997 Elsevier Science Inc.

1. INTRODUCTION

The Craig-Sakamoto theorem on the independence of two quadratic forms can be proved in its simplest form by showing that if A and B are symmetric matrices, then

$$|I - \alpha A - \beta B| = |I - \alpha A||I - \beta B| \qquad \forall \alpha, \beta \tag{1.1}$$

if and only if $AB = 0$. Obviously, if $AB = 0$ then (1.1) holds. The more difficult and intriguing result is the converse. There are many papers dealing with a proof that (1.1) implies that $AB = 0$, as well as a variety of extensions. We do not present a history of the subject, which is well described in several other papers (see Driscoll and Gundberg, 1986; Driscoll and Krasnicka, 1995; Ogawa, 1993; and Reid and Driscoll, 1988). Rather, we attempt to provide an "elementary" proof that requires some simple facts about determinants plus a

LINEAR ALGEBRA AND ITS APPLICATIONS 264:217–223 (1997)

655 Avenue of the Americas, New York, NY 10010

0024-3795/97/$17.00
PII S0024-3795(96)00484-3

certain amount of bookkeeping. Of course, the definition of "elementary" is in the eye of the beholder. The germ of this proof is no doubt somewhere in the literature, but it has not been explicated in a form that makes it widely accessible.

We use the notation $C(i_1, \dots, i_k)$ to denote the determinant of a principal submatrix of an $n \times n$ matrix C with rows and columns $i_1, \dots, i_k$. Sums are over all permutations of $\{1, 2, \dots, n\}$. Thus, for example, $\sum x_1 = \sum x_i, \sum x_1 x_2 = \sum_{i<j} x_i x_j$, and so on, so that if $n = 3$,

$$\sum x_1 C(2,3) = x_1 C(2,3) + x_2 C(1,3) + x_3 C(1,2).$$

Without loss in generality we can assume that $A = D_a = \operatorname{diag}(a_1, \dots, a_n)$ is diagonal and replace the negative signs in (1.1) with positive signs. For simplicity of notation, let $x_i \equiv 1 + \alpha a_i$ and $C \equiv \beta B$. It is a well-known result of Sylvester (see, e.g., Aitken, 1944, p. 87) that

$$|D_x + C| = |C| + \sum x_1 C(2, \dots, n) + \sum x_1 x_2 C(3, \dots, n) + \cdots + \sum x_1 x_2 \cdots x_{n-1} C(n) + \prod x_i. \tag{1.2}$$

Using (1.2) in (1.1), we obtain

$$\begin{aligned} &\beta^n |B| + \beta^{n-1} \sum x_1 B(2, \dots, n) + \beta^{n-2} \sum x_1 x_2 B(3, \dots, n) \\ &\quad + \cdots + \beta \sum x_1 x_2 \cdots x_{n-1} b_{nn} \\ &\quad = \left(\prod_1^n x_i\right)\Big(\beta^n |B| + \beta^{n-1} \sum B(2, \dots, n) + \beta^{n-2} \sum B(3, \dots, n) \\ &\qquad + \cdots + \beta \sum b_{nn}\Big), \end{aligned} \tag{1.3}$$

which holds for all α, β.

Rewrite (1.3) as

$$\begin{aligned} &\beta^n |B| \left(\prod_1^n x_i - 1\right) + \beta^{n-1} \sum B(2, \dots, n) \left(\prod_1^n x_i - x_1\right) \\ &\quad + \cdots + \beta \sum b_{nn} \left(\prod_1^n x_i - \prod_1^{n-1} x_i\right) = 0. \end{aligned} \tag{1.4}$$

Because the left-hand side holds for all β, each coefficient must vanish.

The key point in the proof lies in an examination of but two terms:

$$\sum b_{nn}\left(\prod_1^n x_i - \prod_1^{n-1} x_i\right) = 0, \tag{1.5a}$$

and

$$\sum B(n-1, n)\left(\prod_1^n x_i - \prod_1^{n-2} x_i\right) = 0. \tag{1.5b}$$

Note that

$$\prod_1^m x_i = 1 + \alpha \sum a_1 + \alpha^2 \sum a_1 a_2 + \cdots + \alpha^m \prod_1^m a_i.$$

For each $r = 1, 2, \ldots, n-1$, $\prod_1^n x_i - \prod_1^r x_i$ is a polynomial in α of the form $\alpha d_1 + \alpha^2 d_2 + \cdots + \alpha^n d_n$, where the coefficients are functions of $a_1, \ldots, a_n$ and depend on r. For example,

$$\prod_1^n x_i - x_n = (1 + \alpha a_n)\left(\alpha \sum_1^{n-1} a_i + \alpha^2 \sum_1^{n-1} a_i a_j + \cdots + \prod_1^{n-1} a_i\right).$$

Consequently, the left-hand sides of (1.5a) and (1.5b) are polynomials in α and vanish for all α; hence each coefficient must be zero.

We now examine (1.5a, b) by considering different possible scenarios in which each of $a_1, \ldots, a_n$ do or do not vanish. The case $a_1 = \cdots = a_n = 0$ is trivial, for then $D_a B = 0$, which completes the proof.

Suppose that

$$a_1 \neq 0, \ldots, \quad a_r \neq 0, \quad a_{r+1} = \cdots = a_n = 0, \qquad 1 \leqslant r < n. \tag{1.6}$$

Then $x_{r+1} = \cdots = x_n = 1$, and for any subset $\mathscr{B}$ of $\{r+1, \ldots, x_n\}$,

$$\prod_1^r x_i \prod_{j \in \mathscr{B}} x_j = \prod_1^r x_i.$$

Let $\mathscr{A} = \{1, \ldots, r\}$ and $\mathscr{B} = \{r+1, \ldots, n\}$. The coefficients of α^m in (1.5a) vanish for $m \in \mathscr{B}$. For $m = r$ monomials $a_{i_1} a_{i_2} \cdots a_{i_r}$ vanish if any $i_\alpha \in \mathscr{B}$, and the coefficients of $b_{11}, b_{22}, \ldots, b_{rr}$ are each $a_1 a_2 \cdots a_r$, which yields the equation

$$(b_{11} + \cdots + b_{rr}) \prod_1^r a_i = 0;$$

by assumption, $a_i \neq 0$, $i = 1, \ldots, r$, and hence

$$b_{11} + \cdots + b_{rr} = 0. \tag{1.7}$$

In (1.5b) the coefficient of α^r is critical. All terms not involving $a_1, \ldots, a_r$ vanish, and we obtain

$$\left(\sum_{j=r+1}^{n} B(1,j) + \cdots + \sum_{j=r+1}^{n} B(r,j) + \sum_{\substack{i=1 \\ i<j}}^{r} \sum_{j=1}^{r} B(i,j) \right)$$

$$\times (a_1 a_2 + a_1 a_3 + \cdots + a_{r-1} a_r) = 0; \tag{1.8}$$

the term involving the a's does not vanish, so that

$$\sum_{j \in \mathscr{B}} \sum_{i \in \mathscr{A}} B(i,j) + \sum_{j \in \mathscr{A}} \sum_{i \in \mathscr{A}} B(i,j) = 0, \tag{1.9}$$

where $\mathscr{A} = \{1, \ldots, r\}$, $\mathscr{B} = \{r+1, \ldots, n\}$.

The proof is completed using the determinantal theorem proved below. That is, if $a_1 \neq 0, \ldots, a_r \neq 0$ and $a_{r+1} = \cdots = a_n = 0$, then (1.7) and (1.9) hold. The determinantal theorem then shows that

$$B = \begin{pmatrix} 0 & 0 \\ 0 & B_{22} \end{pmatrix}$$

and hence

$$D_a B = \begin{pmatrix} D_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & B \end{pmatrix} = 0,$$

where $D_1 = \operatorname{diag}(a_1, \ldots, a_r)$. ■

2. A CURIOUS DETERMINANTAL RESULT

The following determinantal result is somewhat surprising. Let $B = (b_{ij})$ be an $n \times n$ symmetric matrix partitioned as follows:

$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix},$$

where B_{11} is an $r \times r$ principal submatrix. Let $\mathscr{A} = \{1, 2, \ldots, r\}$, $\mathscr{B} = \{r + 1, \ldots, n\}$. Denote the determinant of a 2×2 principal submatrix by

$$B(i,j) = \begin{vmatrix} b_{ii} & b_{ij} \\ b_{ij} & b_{jj} \end{vmatrix} = b_{ii}b_{jj} - b_{ij}^2,$$

and for simplicity of notation write $b_i \equiv b_{ii}$.

DETERMINANTAL THEOREM. *If* $b_{11} + \cdots + b_{rr} = 0$ *and*

$$\sum_{\substack{i,j \in \mathscr{A} \\ i<j}} B(i,j) + \sum_{j \in \mathscr{B}} \sum_{i \in \mathscr{A}} B(i,j) = 0, \tag{2.1}$$

then

$$B = \begin{pmatrix} 0 & 0 \\ 0 & B_{22} \end{pmatrix}.$$

Proof. Using the fact that $b_1 = -b_2 - \cdots - b_r$ we evaluate each term in (2.1):

$$\begin{aligned}
\sum_{\substack{i,j \in \mathscr{A} \\ i<j}} B(i,j) &= \sum_{j=2}^{r} B(1,j) + \sum_{2 \leqslant i < j \in \mathscr{A}} B(i,j) \\
&= \sum_{j=2}^{r} b_1 b_j - \sum_{j=2}^{r} b_{1j}^2 + \sum_{2 \leqslant i < j \leqslant r} b_i b_j - \sum_{2 \leqslant i < j \leqslant r} b_{ij}^2 \\
&= -\sum_{j=2}^{r} \sum_{i=2}^{r} b_i b_j - \sum_{j=2}^{r} b_{1j}^2 + \sum_{2 \leqslant i < j \leqslant r} b_i b_j - \sum_{2 \leqslant i < j \leqslant r} b_{ij}^2
\end{aligned}$$

$$= -\sum_{i=2}^{r} b_i^2 - \sum_{2 \leqslant i < j \leqslant r} b_i b_j - \sum_{\substack{i,j \in \mathscr{A} \\ i<j}} b_{ij}^2$$

$$= -\frac{1}{2}\left(\sum_{i=2}^{r} b_i\right)^2 - \frac{1}{2}\sum_{i=2}^{r} b_i^2 - \sum_{\substack{i,j \in \mathscr{A} \\ i<j}} b_{ij}^2, \tag{2.2a}$$

$$\sum_{j \in \mathscr{B}} \sum_{i \in \mathscr{A}} B(i,j) = \sum_{j \in \mathscr{B}} B(1,j) + \sum_{j \in \mathscr{B}} \sum_{2 \leqslant i \in \mathscr{A}} B(i,j)$$

$$= \sum_{j=r+1}^{n} b_1 b_j - \sum_{j=r+1}^{n} b_{1j}^2 + \sum_{j=r+1}^{n} \sum_{i=2}^{r} b_i b_j - \sum_{j=r+1}^{n} \sum_{i=2}^{r} b_{ij}^2$$

$$= -\sum_{j=r+1}^{n} \sum_{i=2}^{r} b_i b_j - \sum_{j=r+1}^{n} b_{ij}^2 + \sum_{j=r+1}^{n} \sum_{i=2}^{r} b_i b_j$$

$$- \sum_{j=r+1}^{n} \sum_{i=2}^{r} b_{ij}^2$$

$$= -\sum_{j=r+1}^{n} b_{ij}^2 - \sum_{j=r+1}^{n} \sum_{i=2}^{r} b_{ij}^2$$

$$= -\sum_{j \in \mathscr{B}} \sum_{i \in \mathscr{A}} b_{ij}^2. \tag{2.2b}$$

The right-hand sides of (2.2a) and (2.2b) are each less than or equal to zero, so that if the sum of these two quantities is zero, then each one must be zero. From (2.2b) we have that $B_{12} = 0$. From (2.2a), $b_2 = \cdots = b_r = 0$, so that $b_1 = 0$, which together with

$$\sum_{\substack{i,j \in \mathscr{A} \\ i<j}} b_{ij}^2 = 0$$

is equivalent to $B_{11} = 0$. ∎

REFERENCES

Aitken, A. C. 1944. *Determinants and Matrices*, 3rd ed., Oliver and Boyd, London.

Driscoll, M. F. and Gundberg, W. R., Jr. 1986. A history of the development of Craig's theorem, *Amer. Statist.* 40:65–70.

Driscoll, M. F. and Krasnicka, B. 1995. An accessible proof of Craig's theorem in the general case, *Amer. Statist.* 49:59–62.

Ogawa, J. 1993. A history of the development of Craig-Sakamoto's theorem viewed from Japanese standpoint, *Proc. Ann. Inst. Statist. Math.* 41:47–59.

Reid, J. G. and Driscoll, M. F. (1988). An accessible proof of Craig's theorem in the noncentral case, *Amer. Statist.* 42:139–142.

Received 3 August 1995; final manuscript accepted 18 February 1996